# Emotional Expression in Oral History Narratives: Comparing Results of Automated Verbal and Nonverbal Analyses*

**Khiet P. Truong[1], Gerben J. Westerhof[2], Sanne M.A. Lamers[2], Franciska de Jong[1,3], and Anneke Sools[2]**

1 **Human Media Interaction**
**University of Twente**
**Enschede, The Netherlands**
{k.p.truong,f.m.g.dejong}@utwente.nl
2 **Department of Psychology, Health and Technology**
**University of Twente**
**Enschede, The Netherlands**
{g.j.westerhof,s.m.a.lamers,a.m.sools}@utwente.nl
3 **Erasmus Studio**
**Erasmus University Rotterdam**
**Rotterdam, The Netherlands**

## Abstract

Audiovisual collections of narratives about war-traumas are rich in descriptions of personal and emotional experiences which can be expressed through verbal and nonverbal means. We complement a commonly used verbal analysis with a nonverbal one to study emotional developments in narratives. Using automatic text, vocal, and facial expression analysis we found that verbal emotional expressions do not correspond much to nonverbal ones. This observation may have important implications for the way narratives traditionally are being studied. We aim to understand how different modes of narrative expression relate to each other, and to enrich digital audiovisual interview collections with emotion-oriented tags.

**1998 ACM Subject Classification** I.2.7 Natural Language Processing, H.3.1 Content Analysis and Indexing, H.3.7 Digital Libraries

**Keywords and phrases** narrative psychology, automatic human behavior analysis, automatic content analysis, verbal emotional expression, nonverbal emotional expression

**Digital Object Identifier** 10.4230/OASIcs.CMN.2013.310

## 1 Introduction

In narrative psychology, life stories are generally seen as cognitive devices that people use to attribute meaning to their lives by telling and structuring their life stories [11, 16]. Numerous studies have shown that people who have suffered traumas can improve their physical and mental health by talking or writing about their experiences [6]. The linguistic expression of emotions is believed to serve as the cognitive encoding of analogous emotional experiences that makes it possible to further attribute meaning to them [13]. Resources that psychologists usually use in their analyses are texts about personal experiences, either written down by

* This work was partially supported by FP7/2007-2013 under grant agreement no. 231287 (SSPNet).



Workshop on Computational Models of Narrative 2013.
Editors: Mark A. Finlayson, Bernhard Fisseni, Benedikt Löwe, and Jan Christoph Meister; pp. 310–314

OpenAccess Series in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

the subject or transcribed from the subject's speech by a human transcriber. The style of writing and the words used in these written narratives reflect how people express their emotional experiences, cognitively processing and structuring them. For example, people use more positive emotion words to write about positive events and more negative content words to write about negative events [8]. In a similar vein, the use of pronouns can provide information about how people are cognitively processing the situation: more first-person singular pronouns are used by people who are experiencing physical or emotional pain since they tend to have their attention drawn to themselves [14]. Based on these findings, narrative psychologists consider written words and language (i.e., verbal behaviors) the medium for assessing and understanding the cognitive processes related to emotional expression and meaning construction.

An intriguing question is how verbal expressions in written narratives relate to non-verbal behaviors in the production of narratives. It is known that elements of non-verbal behavior such as the way people speak and their accompanying facial and bodily expressions reflect cognitive and emotional aspects of the narrator's state of mind, e.g., [4, 9]. The frequency of disfluencies (e.g., filled pauses, hesitations) for example is often linked to planning effort and cognitive load during speaking [7]. Consequently, it seems important to take both modalities, verbal and non-verbal expressions, into account when analyzing emotional experiences in narratives. With the increased availability of multimedia content, in particular oral history interview collections, it becomes possible to study both the verbal and non-verbal content of narratives. In our study, we provide a first investigation into how analyses of verbal and non-verbal emotional content in narratives relate to each other.

We apply our analyses to a large collection of audiovisual testimonies on war-related experiences in Croatia's past (CroMe [3]). For the verbal analysis, we use a computerized text analysis tool, the Linguistic Inquiry and Word Count (LIWC) [12] that is commonly employed by psychologists. For the non-verbal analysis, we automatically extract prosodic features from the narrator's speech. In addition, tools for automatic emotion detection in vocal and facial expressions are applied. These verbal and non-verbal measurements are compared to each other through correlation analysis. Our long-term goal is to develop computational models of narrative structure (based on verbal and non-verbal information) that are informed and validated by psychologists to make meaningful browsing and searching through large collections of narratives possible. More fundamental research goals are to understand how different modes of narrative expression relate to each other and how convergence and divergence between the modes is informative for understanding the possibly different cognitive processes and structures involved in verbal and nonverbal expression.

## 2 Material

We use audiovisual recordings of interviews made in the Croatian Memories project [3]. In this project, citizens of Croatia are interviewed about their personal experiences with war and trauma. The project is covering three timeframes: WWII, the Yugoslav period, and the war of the nineties. Guided by the interviewer's questions and responses, the interviewee (while being continuously on camera) tells personal stories about his/her experiences. Hence, these interviews have narrative characteristics and contain emotional episodes. Each interview is approximately between 40 and 70 minutes long. Currently, 50 interviews have been transcribed and translated from Croatian into English. Metadata of the interviewees (e.g., profession, religion) and summaries of the interviews are available, including English (time-aligned) subtitles. A collection of around 400 interviews will be released in the Fall of 2013, and a similar but smaller dataset is being created in Bosnia and Herzegovina.

# 3 Method

Verbal and non-verbal features were extracted from the transcriptions, audio and video. A manual labelling of who is speaking, the interviewer or interviewee, was performed.

## 3.1 Verbal analysis

The verbal analysis was carried out using the LIWC program [12]. This program automatically counts the percentage of words that have linguistic and/or psychological relevance. For this study, we calculated the percentage of positive emotions words and negative emotion words.

## 3.2 Non-verbal analysis

The non-verbal analysis focused on the vocal and facial modality. For the vocal analysis, we automatically measured F0, intensity and pausing behavior with Praat [2] from the interviewee's speech. These measures are known to be related to emotional and cognitive processes in speaking, e.g., [1, 15]. First, speech activity detection was performed (by thresholding intensity) in order to obtain so-called talkspurts. Talkspurts are defined as stretches of continuous speech with minimal durations of 300 ms, bounded by silences with minimal durations of 200 ms. F0 and intensity were extracted each 10 ms, and averaged over each talkspurt. Pauses are defined as within-speaker silences. In addition to these prosodic measurements, we also obtained emotion-oriented labels(and their corresponding confidence scores) from automatic emotion classification tools such as openSMILE [5] for vocal expressions and CERT [10] for facial expressions. Table 1 gives a summary of all features extracted.

**Table 1** Non-verbal measures automatically extracted from speech and face.

| | |
|---|---|
| Praat [2] | mean and standard deviation pitch, mean and standard deviation intensity, mean duration talkspurt, max duration talkspurt, mean duration pause, max duration pause, $\frac{\text{mean duration pause}}{\text{mean duration talkspurt}}$, $\frac{\text{number of pauses}}{\text{mean duration talkspurt}}$ |
| openSMILE [5] | positive emotion, negative emotion, arousal, valence, anger, boredom, disgust, fear, happiness, neutral, sadness, aggressive, cheerful, intoxicated, nervous, tired, interest |
| CERT [10] | positive emotion, negative emotion, smile, agner, contempt, disgust, fear, joy, sad, surprise, neutral |

# 4 Experimental setup

Based on the LIWC analysis, we selected 6 interviews varying in their overall level of emotional expressivity as well as their ratio between positive versus negative emotion words. Each interview was divided into 1-minute segments and analyzed as described in the methods section above. All the verbal and non-verbal measures were averaged over these 1-minute segments. We used Pearson correlations to compute the association of the percentage of positive and negative emotion words with all nonverbal markers over whole interviews.

## 5 Results

Fig. 1 provides the minute-by-minute graphs of some of the verbal and nonverbal markers of one interview. Such figures provide the basis description per interview. In the top panel figure, the fluctuation of positive and negative emotion words is visualized. It can be seen how the interviewee uses more or less emotion words throughout the interview and how the interviewee switches from a predominant use of positive emotion words during the first fifteen minutes to a predominance of negative emotion words during the second thirty minutes to a more mixed level in the remainder of the interview. This increased use of negative emotion words seems to be associated with a lower talkspurt duration which is reflected in a significant moderate correlation of $-0.409, p < 0.005$ (Bonferroni correction was applied). In a similar way, we also computed correlations between the expression of positive and negative emotions words with all other non-verbal indicators. However, in general, most of the correlations were low and not signficant.

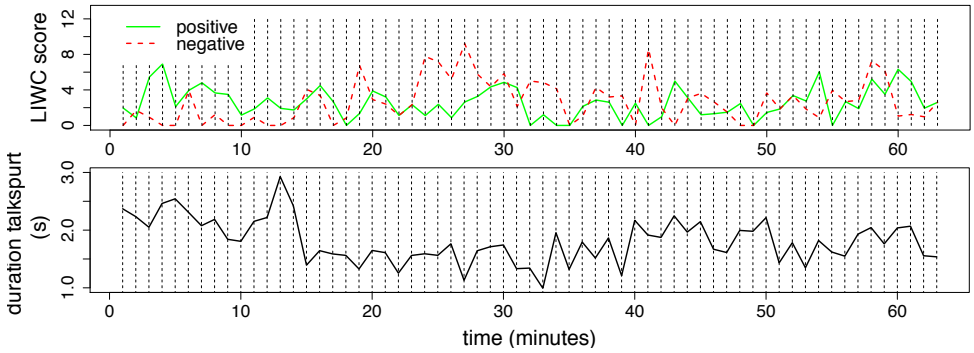


**Figure 1** Examples of verbal and non-verbal measurements of 1 interview. Top pane: LIWC positive and negative emotion scores. Bottom pane: average duration talkspurt (s).

## 6 Discussion

We conclude that the verbal and nonverbal characteristics were unrelated in our study. This might be due to limitations in our analyses. This applies for the verbal analysis, the nonverbal analysis as well as the correlational analysis. In general, we used a one minute interval, whereas longer intervals that are characterized by meaningful story sequences about particular experiences might be more sensitive. Further research is envisaged to check this assumption.

For the verbal analysis, we used the translated transcripts of the originally Croatian narratives. Due to the translation and its purpose for subtitling the interviews, important information on the verbal expression of emotions and its timing might have been lost.

With regard to the nonverbal analysis, we used automatic emotion classification tools that were trained on specific material that not necessarily matches the realistic and noisy audiovisual material we used for our analysis. For example, the facial expression software CERT appeared to work suboptimally as the software is trained on posed facial expressions and requires a frontal face view. In general, we need to analyze more interviews to investigate how idiosyncratic certain speech and facial behaviors are in these specific narratives. In addition, we plan to compare the use of other affect recognition and text mining tools, besides the LIWC.

With regard to the relation between verbal and nonverbal characteristics, we only performed Pearson correlation analysis that does not take the complex nature of the data

into account. In the future, we will use mixed models that can take the dependency and timing within the data of an individual as well as inter-individual differences into account. For these analyses, more interviews need to be analyzed in order to obtain reliable results.

Despite these limitations, the overall lack of relations between the verbal and nonverbal characteristics may have important consequences for theory development as well. It might point to differences in the cognitive processes and structures involved in the verbal and nonverbal expression of emotions. Against the overall lack of relations, the convergence between verbal and nonverbal emotional expression might have a distinct meaning.

## References


1 R. Banse and K. R. Scherer. Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70:614–636, 1996.

2 P. Boersma. Praat: a system for doing phonetics by computer. *Glot International*, 5(9/10):341–345, 2001.

3 Project Croatian memories (CroMe), 2013. `http://balkanvoices.org/?page_id=23`, accessed 24 February 2013.

4 C. Darwin. *The expression of the emotions in man and animals*. D. Appleton and Co., New York, 1872.

5 F. Eyben, M. Wollmer, and B. Schuller. openSMILE – the Munich versatile and fast open-source audio feature extractor. In Alberto Del Bimbo, Shih-Fu Chang, and Arnold W. M. Smeulders, editors, *Proceedings of the 18th International Conference on Multimedia 2010, Firenze, Italy, October 25–29, 2010*, pages 1459–1462, 2010.

6 J. Frattaroli. Experimental disclosure and its moderators: A meta-analysis. *Psychological Bulletin*, 132:823–865, 2006.

7 F. Goldman-Eisler. *Psycholinguistics: Experiments in spontaneous speech*. Academic Press, New York, 1968.

8 J. H. Kahn, R. M. Tobin, A. E. Massey, and J. A. Anderson. Measuring emotional expression with the linguistic inquire and word count. *American Journal of Psychology*, 120:263–286, 2007.

9 M. L. Knapp and J. A. Hall. *Nonverbal Communication in Human Interaction*. Wadsworth Publishing, Boston, MA, USA, 2009.

10 G. Littlewort, J. Whitehill, T. Wu, I. R. Fasel, M. G. Frank, J. R. Movellan, and M. S. Bartlett. The computer expression recognition toolbox (CERT). In *Ninth IEEE International Conference on Automatic Face and Gesture Recognition (FG 2011), Santa Barbara, CA, USA, 21–25 March 2011*, pages 298–305, 2011.

11 D. P. McAdams. Personal narratives and the life story. In O. John, R. Robins, and L. A. Pervin, editors, *Handbook of personality: Theory and research*, pages 241–261. Guilford Press, 2008.

12 J. W. Pennebaker, R. E. Booth, and M. E. Francis. Linguistic inquiry and word count (liwc), 2007. `http://www.LIWC.net`.

13 J. W. Pennebaker and C. K. Chung. Expressive writing and its links to mental and physical health. In H. S. Friedman, editor, *Oxford Handbook of Health Psychology*, pages 417–437, New York, NY, USA, 2011. Oxford University Press.

14 S. Rude, E. M. Gortner, and J. Pennebajer. Language use of depressed and depression-vulnerable college students. *Cognition & Emotion*, 18:1121–1133, 2004.

15 E. E. Shriberg. *Preliminaries to a Theory of Speech Disfluencies*. PhD thesis, University of California at Berkeley, 1994.

16 G. J. Westerhof and E. Bohlmeijer. Life stories and mental health: The role of identifcation processes in theory and interventions. *Narrative Works*, 2(1):106–128, 2012.